Computation of Maximum Likelihood Estimate

Maximize p(d|θ) $(\hat{\theta}_1,\ldots,\hat{\theta}_M) = \arg\max_{\theta_1,\ldots,\theta_M} p(d\,|\,\theta) = \arg\max_{\theta_1,\ldots,\theta_M} \prod_{i=1}^{M} \theta_i^{c(w_i,d)}$

Max. Log-Likelihood $(\hat{\theta}_1,\ldots,\hat{\theta}_M) = \arg\max_{\theta_1,\ldots,\theta_M} \log[p(d\,|\,\theta)] = \arg\max_{\theta_1,\ldots,\theta_M} \sum_{i=1}^{M} c(w_i,d) \log \theta_i$

Subject to constraint: $\sum_{i=1}^{M} \theta_i = 1$

Use Lagrange multiplier approach

Lagrange function: $f(q\,|\,d) = \sum_{i=1}^{M} c(w_i,d) \log q_i + /\,(\sum_{i=1}^{M} q_i - 1)$

$$\frac{\partial f(q\,|\,d)}{\partial q_i} = \frac{c(w_i,d)}{q_i} + / = 0 \quad \rightarrow \quad q_i = -\frac{c(w_i,d)}{/}$$

$$\sum_{i=1}^{M} -\frac{c(w_i,d)}{/} = 1 \rightarrow \quad / = -\sum_{i=1}^{N} c(w_i,d) \rightarrow \quad \hat{q}_i = p(w_i\,|\,\hat{q}) = \frac{c(w_i,d)}{\sum_{i=1}^{M} c(w_i,d)} = \frac{c(w_i,d)}{|d|}$$

Normalized Counts